

SUBSET SELECTION FOR MULTI-GABOR AND NON-ORTHOGONAL WAVELETS EXPANSIONS

L. Rebollo-Neira, J. Fernandez-Rubio and L. Janer[†]

Departament de Teoria del Senyal i Comunicacions, Escola Tecnica Superior
d'Enginyers de Telecomunicació, Campus Nord, UPC,
Edifici D-4, c/. Gran Capita s/n. 08034, Barcelona, Spain
email: laura@gps.tsc.upc.es ; juan@gps.tsc.upc.es

[†] Department de Telecomunicacions i Arquitectura de Computadors
Escola Universitària Politècnica de Mataró (EUPMT)
08303 Mataró, Barcelona, Spain
e-mail: leonard@eupmt.es

ABSTRACT

Non-orthogonal wavelets and Gabor or Multi-windows Gabor expansions [1, 2, 3, 4] involving well-localized synthesis/analysis functions are characterized by being redundant. This entails that the signal modeling is carried out through a rank deficient linear transformation and the expansion coefficients are no unique. In the finite dimensional case one solution for the coefficients (which provides the coefficients of Minimum Norm) is approached by the pseudo-inverse of the concomitant rank deficient transformation. In many applications this makes a great deal of sense. In other applications, however, the model-builder is not interested in a predictor that involves all the redundant factors. Instead, a predictor constructed out of the independent factors is sought. How to pick these factors is a problem of subset selection [5] and we shall advance here a new method for accomplishing such a goal.

1. INTRODUCTION

Algorithms for subset selection should be conceived in order to handle a trade-off between the following matters [5]: i) Estimation of the transformation rank, which is tantamount to estimating an appropriate dimension of the reduced model. ii) Selection of the elements yielding a sub-transformation as well conditioned as possible. iii) Production of a satisfactory residual error.

Our approach is devised so as to simultaneously deal with the above items. We address the problem in an iterative manner. At each step, the element yielding

maximum projection onto the residual vector is selected. This selection allows for the cancellation of such a projection at the next step. We take further care of the reduced model coefficients stability, against different realization of the observation data, by deciding the model dimension tanking into account the imprecision associated with the data.

2. THE PROPOSED APPROACH

2.1. Preliminary considerations

Let us assume that a signal $f^o \in L^2[T_1, T_2]$ is approximated by a function $f \in L^2[T_1, T_2]$ which admits a finite dimensional expansion of the form:

$$f = \sum_{j=1}^J \sum_{m \in \mathcal{Z}_{M_j}} \sum_{n \in \mathcal{Z}_{N_j}} c_{m,n}^j g_{m,n}^j, \quad (1)$$

where J is either the number of different windows (Multi-Gabor case [1, 2, 3, 4]) or mothers wavelet [6]. $\mathcal{Z}_{M_j} \subset \mathcal{Z}$ and $\mathcal{Z}_{N_j} \subset \mathcal{Z}$ are, respectively, finite sets of M_j and N_j integers. The functions $g_{m,n}^j$ are defined as follows:

$$g_{m,n}^j(t) = g^j(t - mb^j) e^{i\omega_o^j nt} \quad ; \quad j = 1, \dots, J$$

(Gabor case)

with b^j and ω_o^j fixed parameters, or

$$g_{m,n}^j(t) = \alpha_o^{j-\frac{m}{2}} g^j(t \alpha_o^{j-m} - nb_o^j) \quad ; \quad j = 1, \dots, J$$

(Wavelets case)

with b_o^j and a_o^j fixed parameters. For each j -value the function $g^j(t)$ represents a given window or mother wavelet.

In order to simplify notation let us re-label indexes by assigning a positive integer k to each triple j, n, m so that $k = 1, \dots, K = \sum_{j=1}^J M_j N_j$ and (1) is recast:

$$f = \hat{A}c = \sum_{k=1}^K g_k c_k \quad (2)$$

where $\hat{A} : l^2(K) \rightarrow L^2[T_1, T_2]$ is a rank deficient operator and c is the vectorial representation of the coefficients $c_k, k = 1, \dots, K$.

Our aim is that of building a predictor $\hat{A}\tilde{c}$ of f^o such that \tilde{c} has at most r nonzero components, with $r \leq \text{Rank}(\hat{A})$. The position of the nonzero entries determines which functions in the model are to be used in approximating the signal f^o . Before addressing the problem of how to pick these functions, let us consider for a while that, out of the set $\{g_k; k = 1, \dots, K\}$, we know a subset of K_o , say, linearly independent functions. Let $\{g_k; k \in K_o\}$ be this subset of functions, K_o being the corresponding subset of K_o indexes. Accordingly, an initial predictor f^{K_o} is built through an operator $\hat{A}_{K_o} : l^2(K_o) \rightarrow L^2[T_1, T_2]$ as:

$$f^{K_o} = \hat{A}_{K_o}\tilde{c} = \sum_{k \in K_o} g_k \tilde{c}_k. \quad (3)$$

Since the functions $\{g_k; k \in K_o\}$ are linearly independent the operator \hat{A}_{K_o} has a unique inverse $\hat{A}_{K_o}^{-1} = (\hat{A}_{K_o}^* \hat{A}_{K_o})^{-1} \hat{A}_{K_o}^*$, where $\hat{A}_{K_o}^*$ indicates the adjoint of \hat{A}_{K_o} . Consequently, the coefficients \tilde{c} are given by $\tilde{c} = (\hat{A}_{K_o}^* \hat{A}_{K_o})^{-1} \hat{A}_{K_o}^* f^o$ and the predictor f^{K_o} is able to be calculated in the fashion [7]:

$$f^{K_o} = \hat{P}_{K_o} f^o = \sum_{n=1}^{K_o} \phi_n \langle \phi_n, f^o \rangle \quad (4)$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product in $L^2[T_1, T_2]$ and the orthonormal functions $\phi_n; n = 1, \dots, K_o$ are obtained by the mapping: $\phi_n = \hat{A}_{K_o} \psi_n / \sqrt{\mu_n}$. The orthonormal vectors $\psi_n \in l^2(K_o)$ are the eigenvectors of the operator $\hat{A}_{K_o}^* \hat{A}_{K_o} : l^2(K_o) \rightarrow l^2(K_o)$, which admits a matrix representation whose elements are: $\hat{A}_{K_o}^* \hat{A}_{K_o}(i, j) = \langle g_i, g_j \rangle; i = 1, \dots, K_o; j = 1, \dots, K_o$. The positive numbers $\mu_n; n = 1, \dots, K_o$ being the corresponding eigenvalues.

2.2. The method

In order to advance the proposed approach, we analyze now the difference $\Delta f^o = f^o - f^{K_o}$ by evaluating all

the projections $\langle g_k, \Delta f^o \rangle; k = 1, \dots, K$. The numbers $\epsilon_k; k = 1, \dots, K$ defined by:

$$\begin{aligned} \epsilon_k = |\langle g_k, \Delta f^o \rangle| &= |\langle g_k, f^o \rangle - \langle g_k, f^{K_o} \rangle| \\ &= |\langle g_k, f^o \rangle - \langle g_k, \hat{P}_{K_o} f^o \rangle| \end{aligned} \quad (5)$$

provide us with a measure for selecting new linearly independent functions. Indeed, if any function g_k belongs to the initial subspace then $\hat{P}_{K_o} g_k = g_k$ and $\epsilon_k = 0$. On the contrary, a large value of ϵ_k implies a large component of Δf^o in the k -direction. According to this feature, an iterative algorithm, whose operational steps we summarize below, naturally evolves.

- i) Consider $K_o = 1$ and pick an arbitrary function, g_{l_1} , say, to build the matrix representation of $\hat{A}_{K_o}^* \hat{A}_{K_o}$ as:

$$\hat{A}_{K_o}^* \hat{A}_{K_o}(1, 1) = \langle g_{l_1}, g_{l_1} \rangle.$$

The single eigenfunction ϕ_1 is then trivial: $\phi_1 = g_{l_1} / \langle g_{l_1}, g_{l_1} \rangle$

- ii) Evaluate $\epsilon_k; k=1, \dots, K$ as

$$\epsilon_k = |\langle g_k, f^o \rangle - \langle g_k, \phi_1 \rangle \langle \phi_1, f^o \rangle|$$

and set l_2 equal to the index k for which ϵ_k is maximum.

Select g_{l_2} as a model function.

- iii) Increase K_o to $K_o + 1$ and build the matrix representation of $\hat{A}_{K_o}^* \hat{A}_{K_o}$ as

$$\begin{aligned} \hat{A}_{K_o}^* \hat{A}_{K_o}(i, j) &= \langle g_{l_i}, g_{l_j} \rangle \quad ; \quad i = 1, \dots, K_o \\ &\quad ; \quad j = 1, \dots, K_o \end{aligned}$$

Compute the eigenvectors $\psi_n; n = 1, \dots, K_o$ and eigenvalues $\mu_n; n = 1, \dots, K_o$ of the above matrix so as to obtain the functions

$$\phi_n = \hat{A}_{K_o} \psi_n / \sqrt{\mu_n} \quad ; \quad n = 1, \dots, K_o$$

- iv) Evaluate the new $\epsilon_k; k = 1, \dots, K$ as

$$\epsilon_k = |\langle g_k, f^o \rangle - \sum_{n=1}^{K_o} \langle g_k, \phi_n \rangle \langle \phi_n, f^o \rangle|$$

and set l_{K_o+1} equal to the index k yielding maximum ϵ_k .

Pick $g_{l_{K_o+1}}$ as a model function.

- v) Compute the residual error

$$\begin{aligned} \|\Delta f^o\|^2 &= \langle \Delta f^o, \Delta f^o \rangle = \|f^o - f^{K_o}\|^2 \\ &= \|f^o - \sum_{n=1}^{K_o} \phi_n \langle \phi_n, f^o \rangle\|^2. \end{aligned}$$

vi) Repeat steps iii), iv), v).

The algorithm is to be stopped when

$$\|\Delta f^o\|^2 \leq \delta,$$

where δ is the precision within which one wishes to approximate the signal f^o . Note: if f^o is a noisy signal, provided that the noise variance σ^2 is known, the signal can be denoised by setting $\delta = \sigma^2$.

The sought vector \tilde{c} is obtained through $\hat{A}_{K_o}^{-1}$, i.e:

$$\tilde{c} = \sum_{n=1}^{K_o} \psi_n \frac{1}{\sqrt{\mu_n}} \langle \phi_n, f^o \rangle. \quad (6)$$

The K_o components of \tilde{c} are the sought nonzero coefficients, c_k , at the positions labeled by the indexes l_k ; $k = 1, \dots, K_o$ that the above algorithm has selected.

3. NUMERICAL EXAMPLE

The purpose of the simple numerical simulation we present here is to illustrate the considerable reduction in the cardinality of the nonzero coefficients (against the Minimum Norm coefficients) achieved by the proposed approach. To this end, we shall deal with the single mother wavelet case corresponding to set $J = 1$ (cf. Eq. (1)).

Consider that in order to represent the chirp signal

$$f(t) = \sin(2\pi t + 20\pi t^2) \quad ; \quad t \in [0, 1]$$

through a discrete set of numbers we use the rather arbitrary model:

$$f(t) = \sum_{m=-5}^0 \sum_{n=0}^{a_0^{-m} b_0^{-1}} c_{m,n} \frac{a_0^{-m}}{p_{m,n}} g(a_0^{-m} t - n b_0), \quad (7)$$

with $a_0 = 2$, $b_0 = 0.5$ and $g(t)$ a mother wavelet as the one proposed in [6], pp 79 (see Fig 1). The numbers $p_{m,n}$ allow for the wavelets normalization in the $[0, 1]$ interval.

The observed data are simulated as

$$f^o(t_i) = f(t_i) + e_i \quad ; \quad t_i = \Delta t(i-1) \quad ; \quad i = 1, \dots, 200$$

where $f(t_i)$ are the sampling values of (7) and e_i are random Gaussian identically distributed variables with mean zero and variance $\sigma^2 = 0.1$. The so simulated data are plotted in Fig 2.

The continuous line of Fig 3 represents the unnoisy

signal we are modeling whereas the dotted line of the same figure depicts the predictions obtained through 35 nonzero coefficients selected by the proposed approach. Fig 4 plots the predicted signal that one would obtain by using, out of the Minimum Norm solution, the 35 coefficients of larger absolute value.

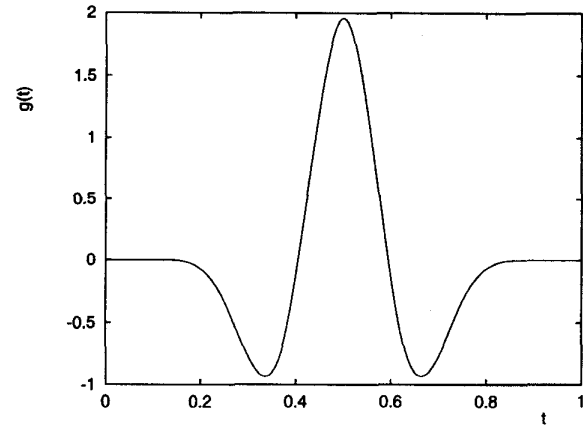


Figure 1: Mother Wavelet

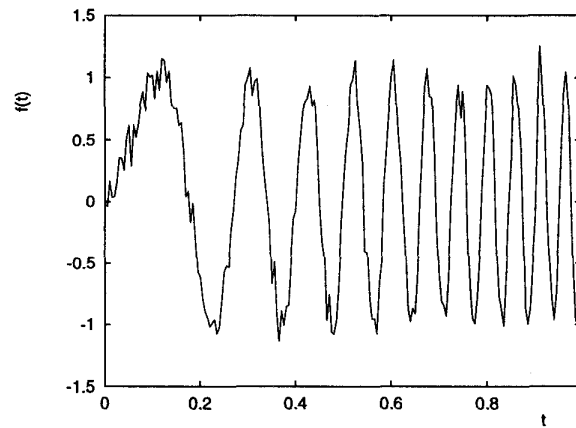


Figure 2: Input noisy data

4. CONCLUSIONS

A method for subset selection has been evolved. The algorithm was devised so as to handle together the following matters: i) Selection of a subset yielding a well conditioned inverse problem while producing an acceptable residual error. ii) Estimation of the reduced model dimension. Both goals have been achieved by iteratively selecting, at each step, the element holding maximum projection onto the previous step residual

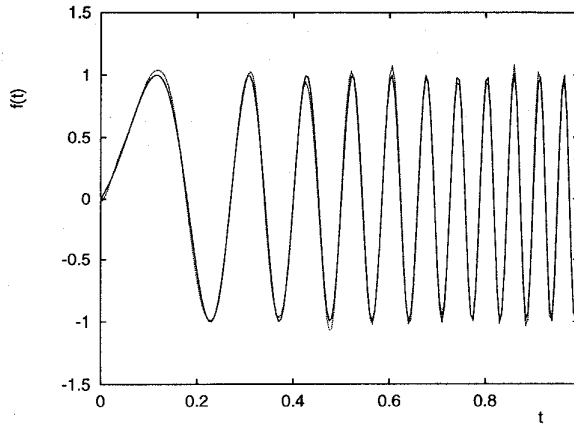


Figure 3: The continuous curve plots the original signal. The dotted line plots the reconstruction obtained through 35 coefficients selected by the proposed approach.

vector. The stability of the model coefficients, against perturbations of the data, has been further considered by deciding the model dimension attending to the data errors. Thus, the dimension of the reduced model depends both on the data errors and the arbitrary function g_{11} that initializes the algorithm. However, regardless of the initial function, the maximum dimension is guaranteed to be at most equal to the numerical rank of the original redundant model.

Acknowledgments

L. Rebollo-Neira is a member of the Scientific Staff of CICPBA (Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina). This work has been supported by CICYT of Spain (TIC95-1022-C05-1, TIC96-0500-C10-01) and CIRIT of Catalunya (1996 GR-00096).

5. REFERENCES

- [1] A. Akan and L. F. Chaporro, "Multi-window Gabor expansion for evolutionary spectral analysis", *Signal Processing*, Vol 63, No 3, 249-262, (1997).
- [2] S. Li "A fast and parametric algorithm for discrete Gabor expansions and the role of various dual windows". *Proc. SPIE'94 Orlando, Conference on Wavelet Applications*, (1995).

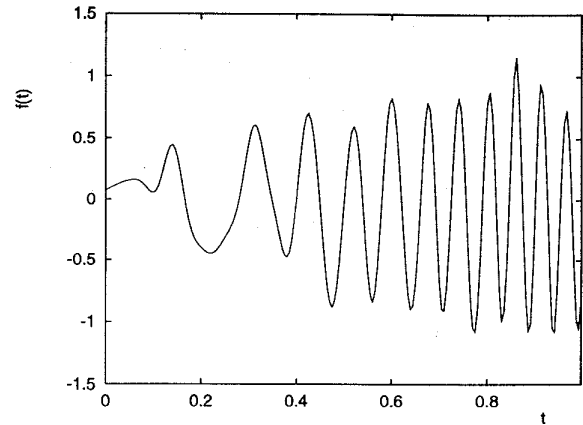


Figure 4: Reconstruction obtained by using, out of the Minimum Norm solution, the 35 coefficients of larger absolute value.

- [3] M. Zibulski, Y. Y. Zeevi, Analysis of multi-window Gabor-type schemes by frame methods. *Proc. SPIE*, San Diego, (1995).
- [4] M. Zibulski, Y. Y. Zeevi, "Discrete Multi-windows Gabor-Type Transforms", *IEEE Transactions on Signal Processing*, Vol. 45, No 6, 1428-1442, (1997).
- [5] G. H. Golub, C. F. Van Loan, "Matrix Computations", The Johns Hopkins University Press, Baltimore, (1989).
- [6] I. Daubechies, "Ten Lectures on Wavelets" CBMS-NSF, SIAM, Philadelphia, (1992).
- [7] L. Rebollo-Neira, A. G Constantinides, T. Stathaki, "Signal representation for compression and noise reduction through frame-based wavelets", *IEEE Transactions on Signal Processing*, Vol.46, No.3, pp.587-597, (1998).